

商品認知と購買意欲向上に 寄与する音の特徴量

テーマ

今後、TVによるターゲティングが可能になると考えられる。また、スマートフォンに視覚を奪われていることによって、TVCM制作においてより聴覚的なアプローチが要求されると考えられる。そこで広告効果の向上に寄与する「調節可能な音の要素」について分析し、また音の調整による広告効果の向上の可能性を確認する。

分析結果

- ・商品認知(PS), 購買意欲(PI)向上に寄与している音の特徴量が分かった
- ・音の特徴量の調整によって、PS, PIの向上が見られた
- ・広告における音の重要性も明らかとなった

考えられる施策

TVerなどの動画配信サービスにおけるCMにおいて、企業が各個人の情報に対応した音の要素の調整をすることによって消費者の聴覚に対して効果的なアプローチができるようにすること。

- 研究背景 4
- 研究方針
 - 1. 目的 5
 - 2. 分析概要 6
 - 3. 使用データ 7,8,9
- 基礎分析
 - 1. CM取得・音の特徴量取得 10
 - 2. 対応分析 11
- 本分析
 - 1. モデルの作成、パラメータのチューニング 12
 - 2. 本分析- 2.SHAP 値によるモデル解釈 13,14
 - 3. 分析 調整する音の特徴量 15
- 結果の解釈 16
- まとめ、今後の課題・展望 17
- 参考文献 18

スマートフォンの使用

<https://gallery.intage.co.jp/tv-sns2/>

テレビとSNSの同時利用者 割合と変化

	土日祝日				平日				コロナ前後 差分			
	2019年 1月	2022年 1月	コロナ前後 差分	2019年 2月	2022年 2月	コロナ前後 差分	2019年 1月	2022年 1月		コロナ前後 差分		
全体	55.3	60.6	+5.3	47.0	57.1	+10.0	53.2	58.3	+5.1	51.9	58.0	+6.2

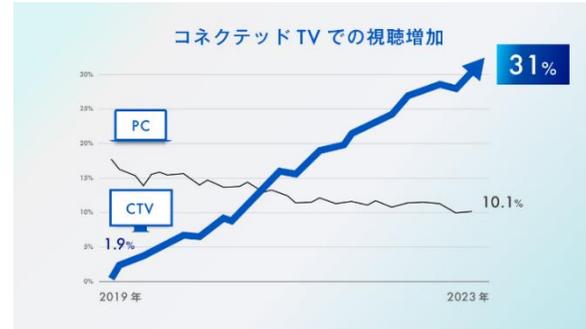
データ：インテージ トSPP
 ベース：テレビとスマホを保有する15歳～69歳男女
 ※各月に1回でも同時利用のログがあった人の割合
 SNSアプリ：Facebook、Instagram、LINE、TikTok、Twitter

- コロナショックの時期と重なって、テレビを見ながらSNS等スマートフォンを利用する人が増加している
- 視覚がテレビから離れている

CMの視覚情報のインパクト
低減する可能性がある

コネクテッドTV

<https://tver.co.jp/news/20230221-1.html/>



- コネクテッドTVが普及している
- 2023年1月にはTVerの視聴媒体の31%がコネクテッドTVとなり、初めて3割を超えた

テレビでも
ターゲティング可能になる

TVer

<https://tver.co.jp/news/20230221-1.html/>



ユーザー数は2022年10月に2,317万MUB(※)、同年12月には2,528万、2023年1月は2,708万と増加し続けている

※MUB : Monthly Unique Browsers

TVerのMUBは
増えていくと予想される

TVerの拡大により可能となるテレビでのターゲティングと、CMの聴覚情報の重要性が増す

本分析は、商品認知や購買意欲を向上させたTVCMの「調整可能であると考えられる音の要素」について分析し、音声調整が広告効果の向上に貢献するのか、また取得可能な個人の情報に対応した音声の調節をどのように最適化すべきかを明らかにすることが目的である。

現在

コネクテッドTV
ユーザーの増加

スマートフォンに
視界を奪われている

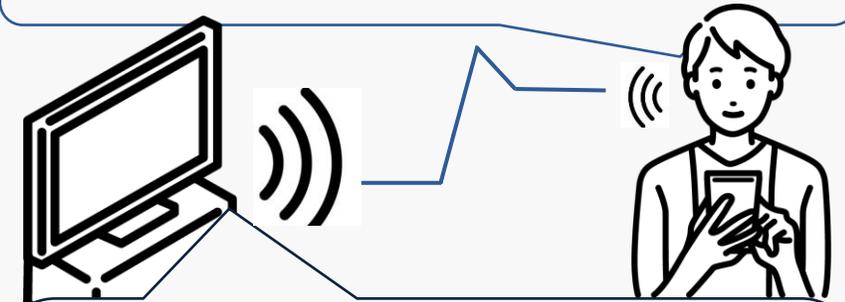
未来

TVCMでの
ターゲティング

聴覚的なアプローチの
重要度の高まり

最適な音声調整

東京都在住、男性、20代…等



BGM有 ・ BGM無
男ナレーション ・ 女ナレーション
db大 ・ db中 ・ db小
f0大 ・ f0中 ・ f0小
⋮

説明変数

- 性別
 - ・年代
 - ・居住地
 - ・趣味
 - ・TVerの利用頻度
 - ・TVerの利用時間
 - ・アンケートデータのRECスケールと商品の対応分析によって得られる値
- ・TVerの利用時間
 - ・音の特徴量
(詳しくは、基礎分析1参照)

**取得可能性の高いと考えられる、
個人に関するデータ**

目的変数

- 商品認知の向上有無(以下PS)
 - 購買意欲の向上有無(以下PI)
- ・ 2時点で行われたアンケート項目の各商品における商品認知(PS)と購買意欲(PI)を加工したもの
 - ・ 具体的には2時点の内、後者のアンケート結果を前者から引き、0以下を0に、それ以外を1に設定

1

アンケート回答者ごとに説明変数を入れたモデルでPS、PIを予測

2

ベースラインと比較して増加した人たちを特定

3

Shap値を求めて解釈する

提供データ1 アンケートデータ



- サンプルサイズ2500の
シングルソースデータ
- 個人属性
- Tverの利用実態
- 商品別の商品認知度
- 商品別の購買意欲

収集データ1 TVCMデータ



CMの放送時期に基づく動画を
公式YouTubeからダウンロード

提供データ2 出稿データ



テレビ番組別の
出稿されたTVCM情報

収集データ2 音声ソフトによる 加工データ



音声ソフト

- Praat (音響分析ソフト)
- Vocal Remover (音声編集サイト)
- librosa (pythonライブラリ)

SEX_CD	性別
AGE	年齢
PREF_CD	居住地

HOB_01_MA	趣味 (スポーツ、フィットネス)
HOB_02_MA	趣味 (ゴルフ)
~	~
HOB_31_MA	趣味 (その他)
HOB_32_MA	趣味 (特にない)

NET_Freq_10_MX	利用頻度 (TVer)
NET_Hour_08_MX	1日当たりの利用時間 (TVer)

NET_Device_08_01_MX	利用デバイス (TVer: 携帯電話・スマートフォン)
NET_Device_08_02_MX	利用デバイス (TVer: パソコン)
NET_Device_08_03_MX	利用デバイス (TVer: タブレット)
NET_Device_08_04_MX	利用デバイス (TVer: その他)

NET_Time_08_01_MX	利用時間帯 (TVer: 朝起きてすぐ)
NET_Time_08_02_MX	利用時間帯 (TVer: 通勤・通学時)
NET_Time_08_03_MX	利用時間帯 (TVer: 昼休み・休み時間)
NET_Time_08_04_MX	利用時間帯 (TVer: 帰宅時)
NET_Time_08_05_MX	利用時間帯 (TVer: 食事中)
NET_Time_08_06_MX	利用時間帯 (TVer: 夜~就寝前)
NET_Time_08_07_MX	利用時間帯 (TVer: 朝起きてすぐ)
NET_Time_08_08_MX	利用時間帯 (TVer: 日中)
NET_Time_08_09_MX	利用時間帯 (TVer: 夜~就寝前)

広告効果を測るための情報

動画配信サービスの具体例として、TVerで取得可能だと考えられる情報

対応分析に使用するRECスケールと商品名

PI_〇〇	〇〇: あなたは今後〇〇を~たいと思いますか。
PS_〇〇	〇〇: あなたは〇〇を知っていますか。
...	...

REC_Scale_01_MX	RECスケール (買うときにはよくバーゲンセールを利用する)
REC_Scale_02_MX	RECスケール (流行中のものを買う)
REC_Scale_03_MX	RECスケール (どの店で買えば得かに行く前によく調べてみる)
REC_Scale_04_MX	RECスケール (そのもののムードや情緒をとくに重視して買う)
REC_Scale_05_MX	RECスケール (買うのは必要最低限にとどめておく)
REC_Scale_06_MX	RECスケール (買うときには店員がすすめるものにする)
REC_Scale_07_MX	RECスケール (買うときにはよく広告しているブランドを買う)
REC_Scale_08_MX	RECスケール (実用性や使いやすさをとくに重視して買う)
REC_Scale_09_MX	RECスケール (見た感じや使いやすさをとくに重視して買う)
REC_Scale_10_MX	RECスケール (できるだけ多くのものを比較したうえで買う物を決める)
REC_Scale_11_MX	RECスケール (新しい物が出たときは人よりも早く買う)
REC_Scale_12_MX	RECスケール (とにかく安くて経済的なものを買う)

item_name	アイテム
-----------	------

取得したCMデータ

アンケート期間中出稿されており、アンケートデータ内の個人が1回以上を見ていて、企業公式アカウントのYoutube上に公開されているCMのデータを用いた。（15秒CM23本と30秒CM11本、計34本のCM）

Google Pixel	アレグラ FX	スーパードライ	スペーシア	アート引越センター	いい部屋ネット
N BOX	ストナリニ S	タウンワーク	薬用 雪肌精	NURO 光	こどもちゃれんじ
PayPay	ビオレ ザ フェイス	タント	iPhone 14 iPhone 14 Pro	アレジオン20	じゃらん
dカード GOLD dカード	ハスラー	マキアージュ ドラマ ティックスキンセンサー ベース NEO	三菱冷蔵庫	めちゃコミック	ソフトバンク デビュー割
ヤクルトシリーズ	マイナビバイト	インディード	進研ゼミ	家庭教師のトライ	コミックシーモア
ドコモ青春割	一番搾り	Galaxy Z Flip シリーズ	クラリチンEX		

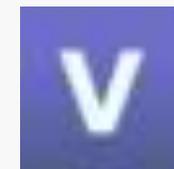
音の特徴量一覧*1

bgm	bgmの有無	cpmn_f	女性による企業名ナレーションの有無
original	オリジナル楽曲の有無	nar_m	男性によるナレーションの有無
vocal	音楽内のヴォーカルの有無	nar_f	女性によるナレーションの有無
vocal_m	音楽内の男性ヴォーカルの有無	ZCR(1ms)	1msあたりに0を通過する回数
vocal_f	音楽内の女性ヴォーカルの有無	mu_f0_mean	音楽のみのf0値の平均
item_song	商品専用音楽の有無	vo_f0_mean	ナレーション部分のf0値の平均
item_nar	商品名のナレーションの有無	muvo_f0_mean	音楽とナレーションのf0値の平均
item_m	男性による商品名のナレーションの有無	unvoice(%)	CM全体における無声時間の割合
item_f	女性による商品名のナレーションの有無	mu_db_mean	音楽部分のdbの平均
cpmn_song	企業専用音楽の有無	vo_db_mean	ナレーション部分のdbの平均
cpmn_nar	企業名ナレーション有無	muvo_db_mean	音楽とナレーションのdbの平均
cpmn_m	男性による企業名ナレーションの有無		

*1 分析対象の34商品のうち、収集できた15秒CM23本と30秒CM11本について、これらの変数を作成した
変数の意味が「有無」となっている変数は、1：有、0：無で定義した質的データである。

アンケート期間中にCMを投稿し、YouTube上で入手可能な34種類の商品に関する変数を取得した

- CMの音源をVocal Remover用いて音楽とナレーション部分に分けた
<https://vocalremover.org/>
- ZCR(1ms)はPythonのlibrosaを利用
<https://librosa.org/>
- f0, db, unvoiced(%)についてはPraatを利用
<https://www.fon.hum.uva.nl/praat/>



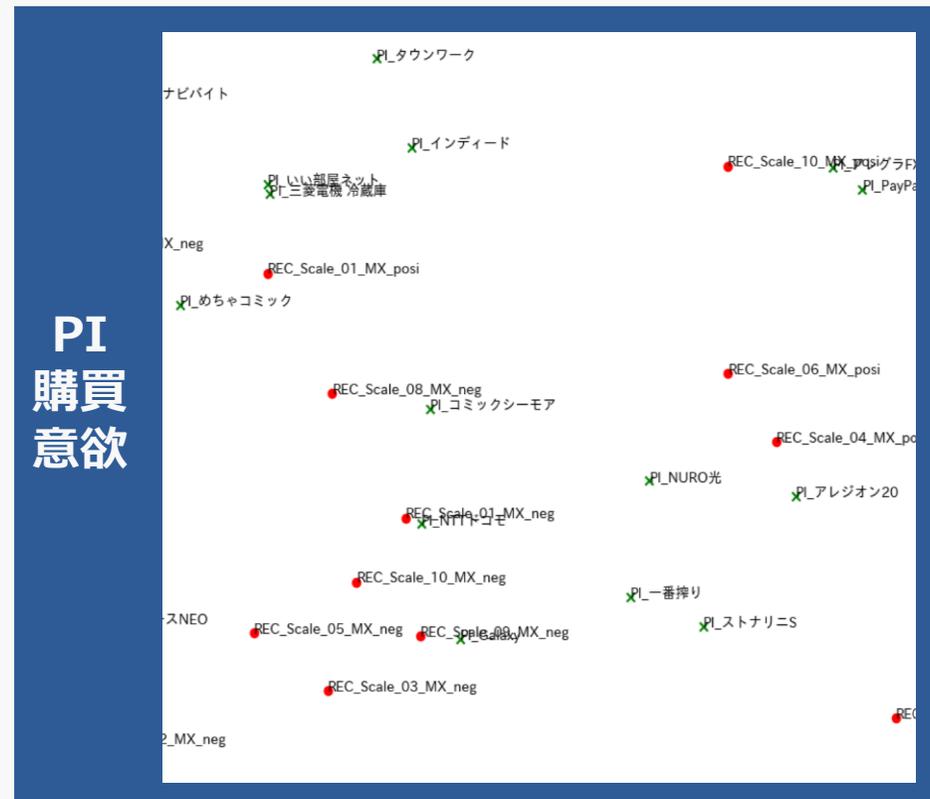
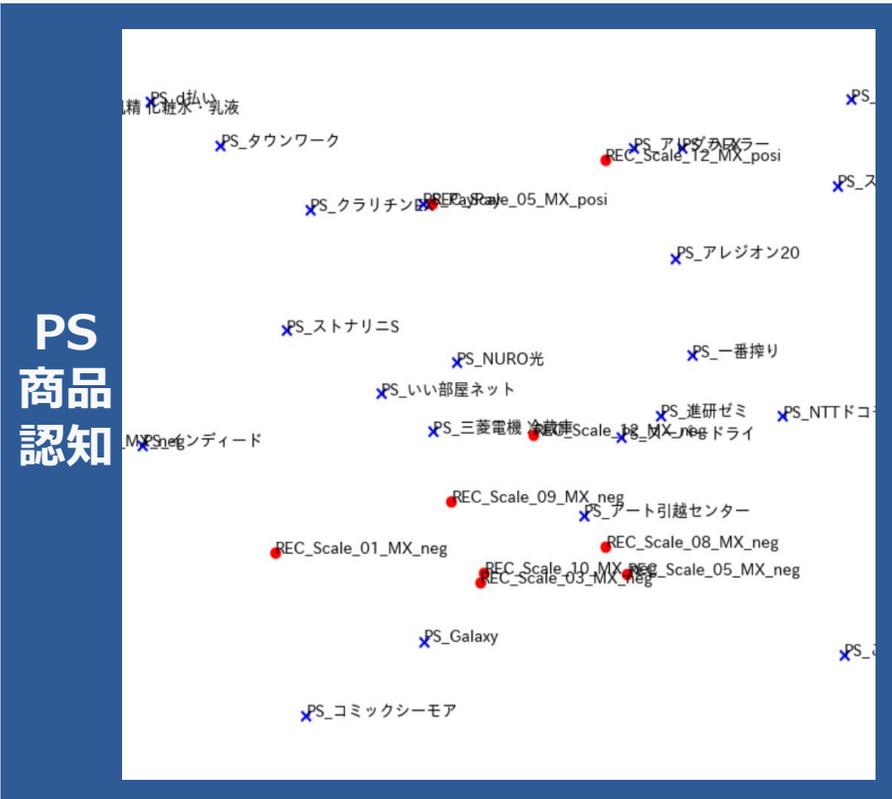
mu : 音源の音楽 **vo : 音源のナレーション**
f0 : 音高の指標 **db : 騒音の指標**
ZCR : ノイジーさ指標

※ZCRはまず音声信号をフレームに分割し各フレーム内で、信号がゼロを横切る回数を数え、フレームの時間の長さで割ることで単位時間あたりのゼロクロスを計算し、これにより音声のやかましさを評価できる。今回は1msあたりを求めた。
また、今回はlibrosaのデフォルト値に従いフレーム数は2048とした。

基礎分析 2 – 対応分析

- CMの広告効果と商品の種類は関連性が高いと考えられるため、商品カテゴリを考慮すべきである
- しかし商品の種類によるカテゴリズでは一種類のカテゴリができてしまったため、対応分析により商品カテゴリに対応するような変数 (PS_1・2, PI_1・2) を取得

(https://corvus-window.com/python_correspondence-analysis/)



使用変数	変数名
REC スケール	REC_Scale_〇〇_MX
CM出稿商品ごとの PS / PI	PS_〇〇 / PI_〇〇

対応分析 (コレスポネンス分析)

質的変数の関係性を視覚化するための多変量解析手法。変数間の関係性を明らかにし、高次元データを低次元に射影
 → どの変数が互いに関連・独立しているのか把握可能

・以下の方法でモデルを作成

推定方法	LightGBM
説明変数	CMデータ デモグラフィック情報 (n=59419,p=82)
目的変数	広告効果量
パラメータチューニング	optuna(auc最大化)
商品認知 : accuracy	0.698
購買以降 : accuracy	0.626

LightGBMによる勾配ブースティング

- ・ 決定木アルゴリズムに基づいた勾配ブースティングのフレームワーク
- ・ 複数のモデルを組み合わせて学習するアンサンブル学習の1つ
- ・ 前の学習機の誤差を新しい学習機が学習して誤差を小さくする
→特徴量分析をするのに必要な
高い精度のモデルを作成することが可能

optuna

- ・ オープンソースのハイパーパラメータ自動最適化フレームワーク
- ・ ハイパーパラメータの値に関する試行錯誤を自動化する
→優れた性能を発揮するハイパーパラメータを
自動的に発見することが可能

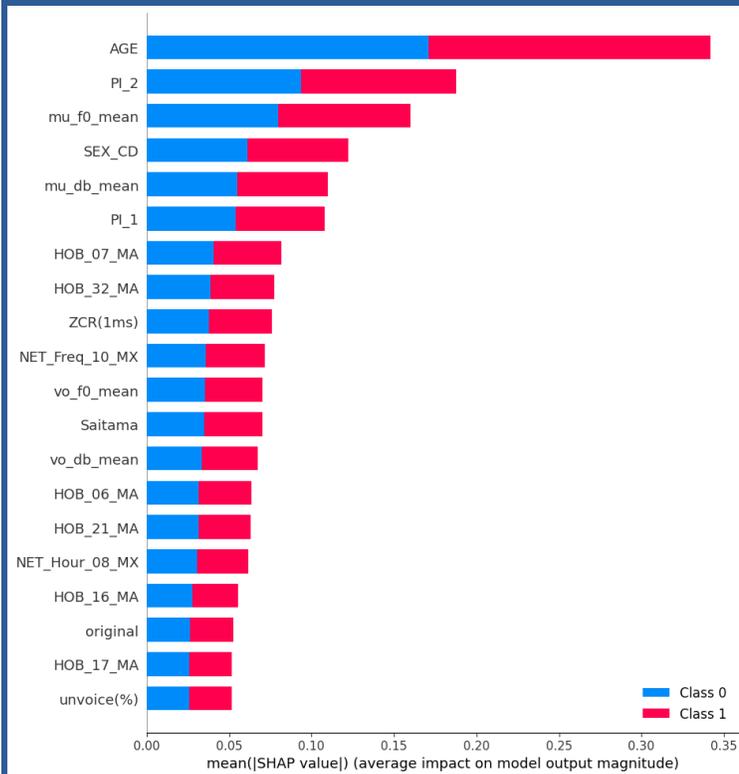
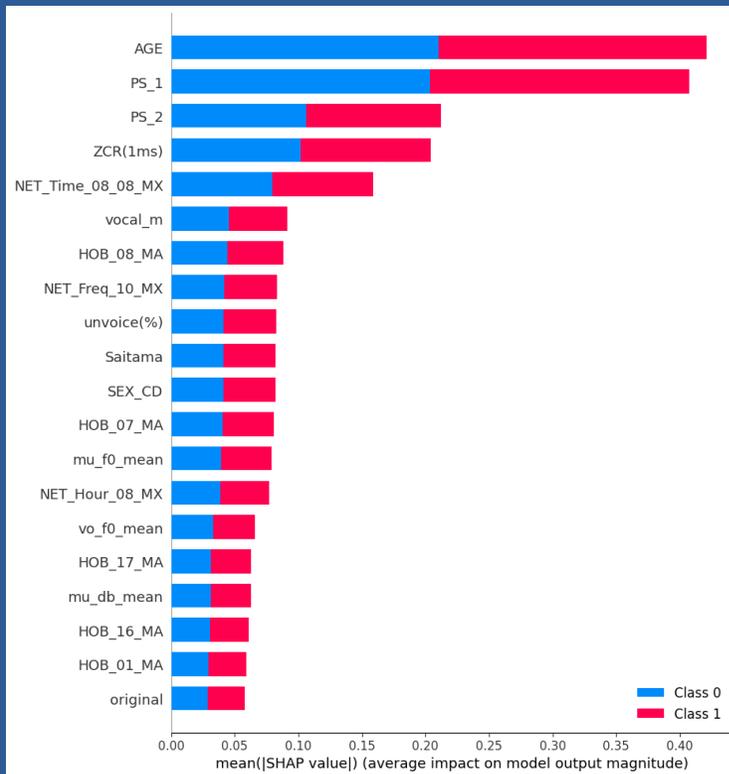
(<https://www.preferred.jp/ja/projects/optuna/>)

Shap はモデルの予測結果に対する各特徴量の寄与度を求めるための手法で、

寄与度として協力ゲーム理論のShapley Valueを用いる (<https://fukufuku3542.hatenablog.com/entry/2020/08/07/165853>)

PS 変数重要度

PI 変数重要度



- 対応分析で取得できたPS_1・2, PI_1・2が上位に表示されている
→ これら二つの変数は精度向上に貢献していると言える
- 音楽のみのf0値の平均(mu_f0_mean)や音楽内の男性ヴォーカルの有無(vocal_m), 1msあたりに0を通過する回(ZCR(1ms))といった、調整可能であると考えられる音の特徴量として設定した変数の重要度が上位に来ている
→ 広告効果向上において、音の調整は有効であることが示唆された

調整可能な音の特徴量として取り上げた変数のうち、重要度が高い順に並べた際の上位3つを実際に調整し、広告効果が向上するか検証する

調整する音の特徴量

	変数名	範囲
商品認知	ZCR(1ms) (1msあたりに0を通過する回数)	1.4 ~ 5.1, 0.1間隔
	vocal_m (音楽内の男性ヴォーカルの有無)	0, 1
	unvoice(%) (CM全体における無声時間の割合)	20 ~ 70, 5 間隔
購買意欲	mu_db_mean (音楽部分のdbの平均)	55 ~ 76, 1間隔
	mu_f0_mean(Hz) (音楽のみのf0値の平均)	100 ~ 301, 50間隔
	ZCR(1ms) (1msあたりに0を通過する回数)	2 ~ 4.1, 0.1間隔

■ 商品認知

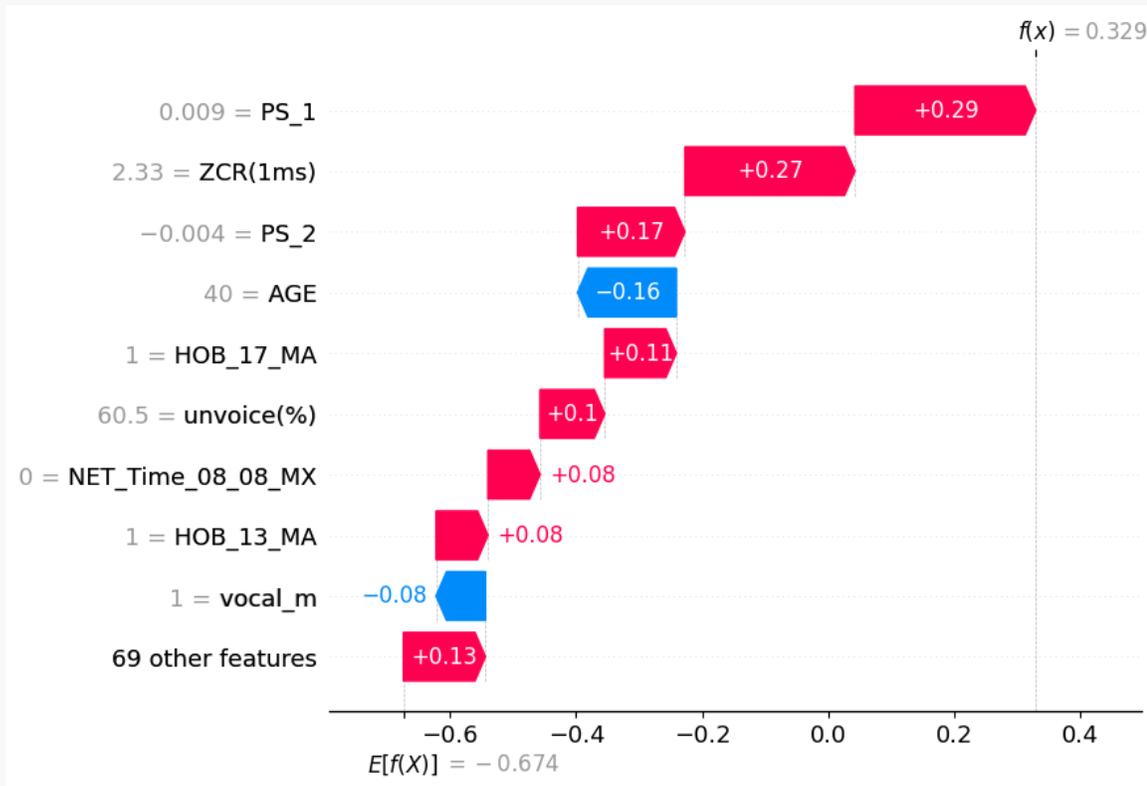
- 調整前：約 14 % → 調整後：約 45 %
- 特徴量の調整によって、各商品に対するサンプルの認知度が増加する割合が全体で **約 31 %** 向上するという結果が得られた

■ 購買意欲

- 調整前：約 18 % → 調整後：約 27 %
- 特徴量の調整によって、各商品に対するサンプルの購買意欲が増加する割合が全体で **約 9 %** 向上するという結果が得られた

■ Waterfall Plot による解釈

パラメータ調整によって、購買意欲の向上がなかった状態から
 向上が生じたサンプルと商品の組み合わせの1つを取り出し、Waterfall Plotを描画



- 対応分析で作成した変数や、今回調整したパラメータであるZCR(1ms)、各種趣味に関する変数が、 $f(x)$ を押し上げている
- $f(x)$ の値が閾値(0.5)未満であることから、元々この組み合わせのPSの値は0と判別されていた
- 音の調整によって、1と判別されるようになったことから、音の調整が $f(x)$ を0.2以上押し上げたことが示唆される

まとめ

- 商品認知・購買意向の向上に貢献している音の特徴量が明らかになった
- 音の特徴量の調整によって、PS及びPIの向上が確認された
- 広告における音の重要性が明らかになった

動画配信サービスへの提言

動画配信サービスは、ユーザーについて取得できるデータを用いて、配信する広告の音の特徴量を最適化し、広告効果が高まる可能性が示唆

今後の展望

- より高い精度のモデルを作るために、さらに多くのCM数が必要であると考えられる
- 音のみ変更した場合のCMのアンケートデータがあれば、各音の特徴量の効果の強さについて言及できると考えられる

・参考文献

- SNSで変わるテレビ視聴 浸透どこまで？「ながら視聴」の実態を読み解く 後編
(<https://gallery.intage.co.jp/tv-sns2/>)
- スマートテレビはどれだけ普及した？ 普及に伴う“テレビの見かた”の変化
(<https://gallery.intage.co.jp/smarttv/>)
- [TVer] 各局ドラマが好調で 過去最高の1月 2,700万 MUB 突破 国内普及率 5割を超えるコネクテッド TV TVer のデバイス別再生割合も3年で15倍以上、初の3割越えに
(<https://tver.co.jp/news/20230221-1.html>)
- テキストアナリティクスの基礎と実践」をpythonで読む3 テキストの特徴分析
(https://qiita.com/tanaka_benkyo/items/58f5ace6d7668688117f)
- SHAPを用いたモデルの解釈
(<https://fukufuku3542.hatenablog.com/entry/2020/08/07/165853>)
- zcr - Zero Crossing Rate - Music Information Retrieval
(<https://musicinformationretrieval.com/zcr.html>)
- Praat: doing phonetics by computer
(<https://www.fon.hum.uva.nl/praat/>)
- Optuna株式会社Preferred Networks Optuna ハイパーパラメータ自動最適化フレームワーク
(<https://www.preferred.jp/ja/projects/optuna/>)
- 対応分析：【Pythonで行う】コレスポンデンス分析（MCA）
(https://corvus-window.com/python_correspondence-analysis/)
- LightGBM : LightGBM | 用語解説 | 野村総合研究所(NRI)
(https://www.nri.com/jp/knowledge/glossary/lst/alphabet/light_gbm)